UDC 551.501.7:519.24:551.507.362.2 NIMBUS 3 IRIS

Application of Stepwise Multiple Regression Techniques to Inversion of Nimbus "IRIS" Observations

GEORGE OHRING [1]—*GCA Corporation, Bedford, Mass.*

ABSTRACT—Stepwise multiple regression analyses are used to explore the statistical (linear regression) relationships between satellite-observed earth–atmosphere emission spectra and meteorological parameters. The stepwise regression technique permits screening of a large number of potentially useful spectral observations (predictors) to isolate those few that contribute most to the explanation of the variance of a particular meteorological parameter. Such a technique is particularly useful when applied to complete spectra of the type obtained by the IRIS (infrared interferometer spectrometer) instrument on the recent Nimbus meteorological satellites. Emphasis is placed upon inferences of key meteorological parameters not usually obtained from routine inversion of satellite spectral observations. The technique is applied to a sample of Nimbus 3 IRIS spectra. The results indicate that information on atmospheric temperatures, geopotential heights of pressure surfaces, tropopause pressure, and tropopause temperature can be obtained directly from the satellite observations with the use of simple linear relationships having only a few terms each. Based upon the results of this exploratory study, suggestions are made for futher development and exploitation of the stepwise regression analysis technique and its application to the problem of inferring meteorological parameters from earth–atmosphere emission spectra of the IRIS type.

1. INTRODUCTION

Statistical inversion techniques for obtaining atmospheric structure information from remote sensing of atmospheric radiation have been discussed by Westwater and Strand (1968), Conrath (1969), and others, and have actually been applied to satellite radiation observations by Smith et al. (1970) and Conrath et al. (1970). In these techniques, multiple regression analysis is performed on a data set of simultaneous radiation and meteorological observations to derive an equation relating the meteorological parameter to the radiation observations. In the present study, we apply a stepwise multiple regression scheme, rather than straight multiple regression, to a sample of Nimbus 3 infrared interferometer spectrometer (IRIS) observations. The stepwise technique is used to develop linear relationships between the meteorological parameters (predictands) and the satellite radiation observations (predictors). The advantage of stepwise regression is twofold.

1. It permits screening of a large number of potential predictors—such as those associated with a complete IRIS spectrum—to obtain the best ones for specifying any meteorological parameter.
2. It can be used to limit the number of predictors to the few that make the most important contributions to explaining a particular meteorological parameter's variance. Theoretical studies (Twomey 1965, 1966) and analyses of weighting functions (Rodgers 1970) show that the information available in the spectrum of the 15-μm CO_2 band (from which the temperature profile is usually derived) is limited to something like four independent pieces. Thus, most of the information on the temperature at any one atmospheric level should be obtainable using only a few predictors.

[1] Now affiliated with the Department of Environmental Sciences, Tel-Aviv University, Ramat-Aviv, Israel

The meteorological parameters usually inferred from satellite-observed infrared spectra are temperature, water vapor amounts, and ozone amounts. We may call these the primary meteorological variables because they, together with clouds if present, control the nature of the observed infrared spectrum. However, other meteorological parameters are related to the primary variables, and hence, information about these parameters should also be contained in the observed spectra. That this is so is evidenced by the relationships obtained between the broadband radiation observations of the TIROS and Nimbus satellites and such meteorological parameters as geopotential height (Jensen et al. 1966), vertical motion and meridional flow (Shenk 1963), areas of precipitation (Chekirda and Iakovleva 1967), stratospheric flow patterns (Zak and Panofsky 1968), and jet stream location and characteristics (Zhvalev et al. 1967, Martin and Salomonson 1970). Thus, in the present study, we have applied the stepwise regression technique not only to the temperature profile inversion problem, but also to the inference of geopotential height of pressure surfaces, tropopause temperature, and tropopause pressure directly from satellite spectra.

2. DATA AND PROCEDURES

To carry out the stepwise regression analyses we need concurrent satellite radiation observations and conventional meteorological observations. The satellite-observed spectra used in this study were measured by the IRIS experiment flown on the Nimbus 3 satellite. The IRIS instrument is described in detail by Hanel et al. (1970) and some of the results obtained with it are discussed by

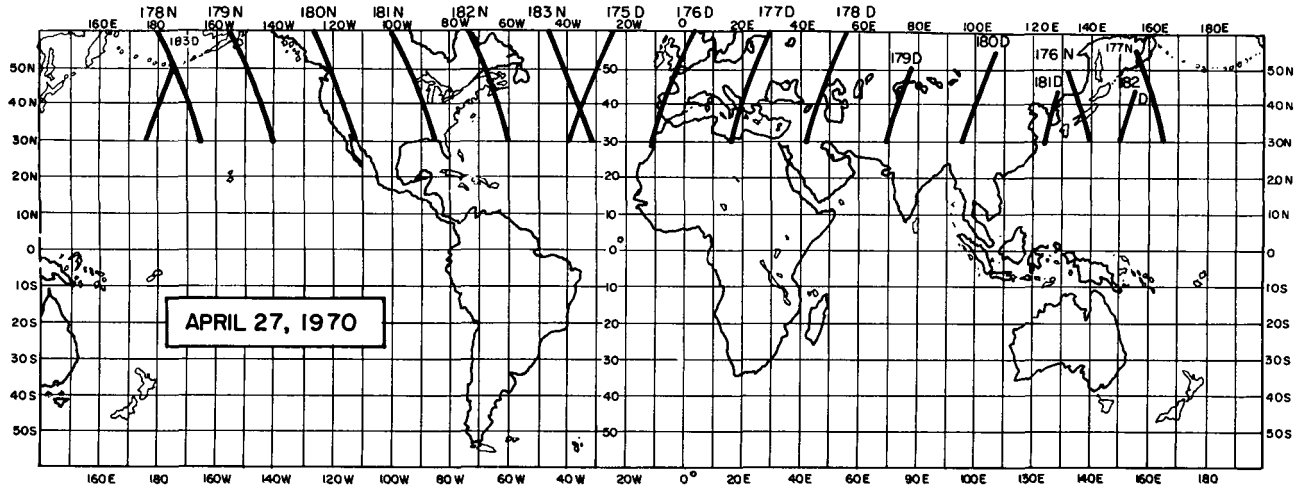

FIGURE 1.—IRIS spectra were obtained for the subsatellite track segments shown on the map. The numbers refer to orbit numbers, D to day, N to night.

Conrath et al. (1970). Briefly, IRIS measures the earth–atmosphere thermal emission spectrum from 400 to 2000 cm^{-1} (5–25μm) with a spectral resolution of 5 cm^{-1} and spacing of about 2.1 cm^{-1}. The measurements are made with a Michelson interferometer that has a radiometric precision and accuracy of about 1 percent. However, after the first 50 orbits, the data from the 1400–2000 cm^{-1} range are unreliable and are not used in the present study. Because of the 8° field of view of the instrument, the radiation that gives rise to a single spectrum originates from a horizontal circular area on earth 150 km in diameter, which is thus the effective horizontal resolution of the observation. Consecutive spectra are of the order of 100 km apart.

The IRIS data were obtained from the Nimbus 3 IRIS archival tape for Apr. 27, 1970. [Information on the format of the IRIS archival data tapes is contained in the *Nimbus 3 User's Guide* (Goddard Space Flight Center 1969)]. On this day, approximately 4,000 spectra were measured. To limit the regression analysis to regions of the globe where reasonable amounts of conventional meteorological data are available, we used only spectra observed between 30° and 60°N in the study. Five hundred spectra fell into this category, and the orbital segments during which these spectra were obtained are shown in figure 1.

Conventional meteorological data for the time period of the satellite observations were obtained from a meteorological data tape supplied by NASA-Goddard Space Flight Center. The data tape contained gridded meteorological data based upon standard U.S. National Weather Service analyses for the Northern Hemisphere. The meteorological data are for two times daily—the standard observing times of 0000 and 1200 GMT.

To obtain concurrency of the satellite observations and the meteorological observations, one must interpolate the meteorological data in both time and horizontal space to the times and geographical locations of the satellite observations. This meshing of the meteorological data with the IRIS observations was accomplished by using linear interpolation of the meteorological data in both time and space.

From the meteorological data set, we selected tropopause pressure, tropopause temperature, and the geopotential heights and temperatures at 850, 700, 500, 300, 200, and 100 mb as predictands. Potential predictors were obtained from the IRIS spectra. In the spectral range 400–1400 cm^{-1}, radiation intensities at about 500 wave numbers are available as potential predictors. With the stepwise regression analysis technique, it is possible to consider each of these intensities as a potential predictor. However, observations at neighboring wave numbers are not independent and, hence, the 500-odd available radiation intensities are not independent observations. In the present study, which is exploratory in nature, the number of potential predictors was limited to 25. Potential predictors were selected from the water vapor, carbon dioxide, and ozone bands, and the atmospheric window. Instead of using the radiation intensities directly, we converted them to brightness temperatures with the use of the Planck law. In addition to these brightness temperatures, two other potential predictors were defined on the basis of the following reasoning. If we consider the tropopause to be the level of lowest temperature, then the wave number within the 15-μm CO_2 band whose weighting function peaks at the tropopause level should record the lowest brightness temperature. As the tropopause moves up and down, different wave numbers will have their weighting function peaks at the tropopause level. Thus, the wave number of minimum brightness temperature should be related to the height of the tropopause. Hence, the wave number, on one side of the 15-μm CO_2 band, having the lowest brightness temperature is a potential predictor of tropopause pressure. This line of reasoning also suggests that the minimum brightness temperature itself should be related to the tropopause temperature. Thus, both the minimum brightness temperature in the 670–720 cm^{-1} range and the wave number at which it occurs are included among the potential predictors. A listing of all 25 potential predictors is contained in table 1.

To perform the stepwise multiple regression analysis, we utilized computer subroutines from the International Business Machines Scientific Subroutine Manual with only minor modifications.[2] The subroutines perform a

[2] Mention of a commercial product does not constitute an endorsement.

TABLE 1.—*List of potential predictors*

Region of spectrum	Predictor number	Predictor [Brightness temperature, T_B, at wave number ν (cm^{-1})]
H_2O rotational band	1	400
	2	450
	3	500
Peak CO_2 absorption	4	668
15-μm CO_2 band	5	670
	6	680
	7	691
	8	702
	9	712
	10	722
	11	733
	12	744
	13	754
	14	764
Atmospheric window	15	900
9.6-μm O_3 band	16	1040
	17	1052
	18	1064
6.3-μm H_2O band	19	1198
	20	1249
	21	1299
	22	1350
	23	1400
Minimum brightness temperature in 15-μm CO_2 band	24	T_B(min)
Wave number of minimum brightness temperature in 15-μm CO_2 band	25	$\nu[T_B(\text{min})]$

stepwise multiple regression analysis for a predictand and a set of potential predictors. In each step of the regression $i=1, 2, 3, \ldots, q$ where q is the number of potential predictors, the abbreviated Doolittle method (Bennett and and Franklin 1954, app. A) is used to calculate which potential predictor results in the largest reduction in the sum of the squares of the deviations at this step and hence should enter the regression. In the present study, we have used a threshold value of 2 percent for the entry of a predictor. That is, the regression is terminated when no potential predictor produces more than a 2-percent reduction in the sum of the squares. This serves to eliminate predictors whose contribution to variance reduction is of the same order as their own observational errors.

Because the infrared spectra are affected by the presence of clouds, regression analyses were also performed on a subset of the complete 500-case data set—the subset consisting of clear-sky cases only. A particular infrared spectrum was classified as a clear-sky case if the absolute difference between the window brightness temperature and the actual surface temperature was less than 5°C. Sixty-three spectra, or 13 percent of the complete data set, fell into this clear-sky subset.

3. RESULTS

Correlation Matrix

Before analyzing the results of the regression analyses, it is of interest to examine the correlations among the various potential predictors. These correlations are contained in the correlation matrix of the predictors that is shown in table 2 and is based upon data from all 500 IRIS spectra. As discussed earlier, the predictors) are primarily brightness temperatures, and the wave numbers shown at the top and side of table 2 indicate the wave numbers to which the brightness temperatures apply. There are two basic reasons for the existence of high correlations among the potential predictors. First, a number of different wave numbers will be sensitive to essentially similar portions of the atmosphere because of the overlap of weighting functions—primarily of one band, but also of several bands. Brightness temperatures at these wave numbers will be highly correlated. Second, brightness temperatures at different wave numbers may be correlated even though they are sensitive to different parts of the atmosphere because of natural interlevel correlations of meteorological parameters. For example, the presence of an inverse relationship between tropospheric and stratospheric temperatures has been known for some time now. The effect of overlapping weighting functions can be seen, for example, by examining the correlation matrix for correlations between the brightness temperatures of any two neighboring wave numbers within the same absorption band. The effect of distant interlevel correlations of atmospheric structure is manifest in the negative correlations between brightness temperatures at the edge (high wave numbers) of the 15-μm CO_2 band (low absorption, hence, sensitive to lower troposphere and surface) and those near the center (low wave numbers) of the band (high absorption, hence, sensitive to stratosphere and upper troposphere). This negative relationship exists even though, of the 500 spectra upon which the correlations are based, most are affected by clouds that tend to hide lower levels of the atmosphere from view and, hence, decrease the magnitudes of these negative correlations. That this is the case can be seen in table 3, which shows, for the complete data set (cloudy plus clear skies) and for clear skies only, correlations between the brightness temperature at 668 cm^{-1} (peak CO_2 absorption; weighting function peak in stratosphere at about 20 km or about 50 mb) and the brightness temperatures at other wave numbers of the CO_2 band. The high negative correlations between the more transparent and the more opaque parts of the 15-μm CO_2 band for clear skies imply that, even in the presence of cloudiness, one should be able to specify something about the meteorological structure of the region below the clouds by using the more opaque channels as predictors. In such cases, the predictability of the lower atmosphere structure will depend not on a direct sensitivity to that layer's infrared emission but solely on the degree of correlation

TABLE 2.—*Correlation matrix of predictors (wave numbers in cm^{-1})*

		H_2O band						CO_2 band							Win-dow	O_3 band			H_2O band					T_B (min)	$\nu[T_B$ (min)]
		400	450	500	668	670	680	691	702	712	722	733	744	754	764	900	1040	1052	1064	1198	1249	1299	1350	1400		
H_2O band	400	1.0	0.74	0.73	0.20	0.18	0.09	−0.06	0.33	0.66	0.73	0.71	0.70	0.67	0.63	0.60	0.57	0.55	0.61	0.63	0.63	0.55	0.57	0.27	−0.05	−0.17
	450		1.0	.99	.06	.03	−.09	−.14	.26	.73	.90	.94	.94	.92	.88	.83	.66	.64	.81	.84	.83	.66	.73	.38	−.13	−.25
	500			1.0	.03	.01	−.12	−.15	.25	.73	.90	.95	.96	.93	.89	.85	.67	.65	.82	.86	.84	.67	.74	.38	−.14	−.26
CO_2 band	668				1.0	.99	.84	.34	.84	.60	.27	.02	−.03	−.13	−.21	−.27	.13	.15	−.09	−.10	−.10	−.03	−.06	−.01	.31	.04
	670					1.0	.87	.37	.84	.57	.24	−.01	−.06	−.16	−.24	−.30	.10	.12	−.13	−.13	−.13	−.05	−.09	−.01	.34	.05
	680						1.0	.65	.65	.32	.04	−.10	−.21	−.26	−.33	−.38	−.15	−.12	−.33	−.29	−.29	−.19	−.22	−.06	.66	.12
	691							1.0	.43	−.11	−.24	−.29	−.28	−.30	−.33	−.35	−.43	−.43	−.47	−.36	−.39	−.33	−.33	−.05	.98	−.02
	702								1.0	.69	.41	.23	.19	.05	−.01	−.09	.24	.23	.08	.08	.06	.09	.04	.09	.38	−.19
	712									1.0	.91	.78	.74	.65	.58	.51	.72	.72	.66	.63	.62	.56	.52	.24	−.13	−.21
	722										1.0	.96	.94	.90	.86	.81	.81	.80	.87	.87	.86	.71	.70	.30	−.23	−.23
	733											1.0	.99	.97	.94	.90	.78	.76	.91	.92	.91	.75	.75	.32	−.28	−.25
	744												1.0	.98	.96	.02	.76	.74	.91	.93	.92	.74	.75	.32	−.27	−.25
	754													1.0	.99	.97	.73	.71	91	.95	.94	.73	.76	.31	−.28	−.23
	764														1.0	.99	.72	.70	.92	.96	.95	.72	.74	.28	−.31	−.23
Window	900															1.0	.71	.68	.91	.95	.94	.70	.72	.26	−.33	−.22
O_3 band	1040																1.0	.98	.91	.76	.78	.76	.66	.30	−.44	−.12
	1052																	1.0	.89	.73	.75	.76	.66	.20	−.43	−.13
	1064																		1.0	.94	.94	.80	.76	.31	−.46	−.21
H_2O band	1198																			1.0	.98	.75	.77	.30	−.35	−.21
	1249																				1.0	.77	.77	.30	−.37	−.21
	1299																					1.0	.76	.44	− 32	−.12
	1350																						1.0	.47	−.31	−.12
	1400																							1.0	−.05	−.15
	T_B (min)																								1.0	.00
	$\nu[T_B$ (min)]																									1.0

TABLE 3.—*Correlations between brightness temperature at 668 cm^{-1} and brightness temperatures at other wave numbers of the 15-μm CO_2 band*

	Wave number (cm^{-1})									
	670	680	691	702	712	722	733	744	754	764
Complete data set	0. 99	0. 84	0. 34	0. 84	0. 60	0. 27	0. 02	−0. 03	−0. 13	−0. 21
Clear skies only	0. 96	0. 90	0. 84	0. 51	−0. 75	−0. 81	−0. 81	−0. 79	−0. 81	−0. 84

between lower and upper atmosphere structure. Such correlations are exploited in the statistical inversion technique developed by Smith et al. (1970).

Temperature

Results of the stepwise multiple regression analysis with temperature as the predictand are contained in table 4. This table shows, for pressure levels from 850 to 100 mb, the coefficient of multiple correlation between the selected predictors and the temperature, the selected predictors, and the constant term (intercept) and regression coefficients of the regression equation. There are several points worthy of note:

1. Despite the presence of cloudiness in most of the cases of the complete data set, correlation coefficients for the lower atmospheric levels are still quite high, indicating that information on temperatures below cloud levels can be obtained from the infrared spectral data. As pointed out in the previous section, this is due to the strong interlevel correlations between lower atmosphere and upper atmosphere temperatures.

2. No more than three predictors are necessary to obtain the correlations shown in table 4, and, in fact, one predictor will suffice for the lower half of the atmosphere. This type of result was anticipated and was one of the reasons for the use of the stepwise multiple regression scheme.

3. The best predictor of lower atmosphere temperatures is T_B(min), the lowest brightness temperature along one side of the 15-μm band. Although initially selected as a potential predictor because of its possible use in specifying tropopause temperature, this predictor has been found to be the best predictor for specifying lower atmosphere temperatures. This finding illustrates one of the advantages of the stepwise technique in combination with a complete infrared spectrum; that is, unanticipated relationships may be uncovered.

4. If T_B(min) had not been available as a predictor, the brightness temperature at 691 cm^{-1} would have been best for specifying the temperature at the three lowest atmospheric levels. The correlation coefficients of T_B(691 cm^{-1}) with atmospheric temperatures at these three levels are slightly less ($\sim$0.01) than those shown in table 4 for T_B(min). This suggests that either T_B(min) or T_B(691 cm^{-1}) could be used with about equal effectiveness to specify temperatures at 500 mb and below. This further suggests that if both are available, both can be used to specify a temperature, and an average of the two can be taken as the final temperature estimate. This procedure reduces the effects of errors in the radiation observations and should lead to a better estimate of temperature than could be obtained using one of the predictors alone.

Results of the regression analyses for temperature in the clear-sky cases are shown in table 5. Except for the

TABLE 4.—*Results of stepwise multiple regression analysis of temperature (°C) using complete data set*

Pressure level (mb)	Coefficient of Multiple correlation	Predictors	Intercept and regression coefficients			
850	0.76	T_B (min)	479	−2.14		
700	.87	T_B (min)	486	−2.21		
500	.90	T_B (min)	452	−2.12		
300	.81	691 702 670	165	−1.26	0.627	−0.313
200	.67	691 $\nu[T_B(\text{min})]$	−50.4	1.18	−0.387	
100	.87	T_B (min) 712	−394	1.74	−0.208	

TABLE 5.—*Results of stepwise multiple regression analysis of temperature (°C) using clear-sky data*

Pressure level (mb)	Coefficient of multiple correlation	Predictors			Intercept and regression coefficients			
850	0.93		764		−255	0.930		
700	.95	1064	702		12.8	0.699	−0.884	
500	.94	1064	702	450	73.4	0.497	−1.24	0.219
300	.78	1249	754		−97.4	0.469	−0.282	
200	.86	702	670	733	−206	3.48	−2.04	−0.653
100	.91	1064	691		−79.5	−0.383	0.553	

300-mb level, all levels now have multiple correlation coefficients greater than 0.85. Once again, these correlations are achieved with three or fewer predictors. Although, as expected, most of the selected predictors are brightness temperatures within the 15-μm CO_2, there are several outside the band. At 700 and 500 mb, T_B(1064 cm^{-1}), in the 9.6-μm O_3 band, is the best single predictor of temperature. However, it is only slightly better than brightness temperatures in the CO_2 band. These results reinforce the recommendation made above; that is, to make use of this redundancy in the spectral observations to improve the temperature specification by averaging temperatures obtained from several regression equations, thus helping to eliminate the effects of random errors in the spectral observations.

A comparison between the coefficients of multiple correlation for the complete data set and for clear skies only is shown in figure 2. For all pressure levels except 300 mb, the coefficient of multiple correlation is, as expected, higher in the case of clear skies. The anomaly at 300 mb presumably represents a quirk in the data sample.

The usefulness of the derived regression equations can best be evaluated by comparing the standard error of estimate of the temperature specification with the standard deviation of temperature in the data sample. If the regression technique represented no skill in the specification of temperature, this would manifest itself by a standard error of estimate equal to the standard deviation. The high coefficients of correlation obtained for temperature indicate that such is not the case here, as is illustrated in figure 3, which compares the standard errors with the standard deviations. For clear skies, the standard error of estimate is between 2° and 3°C at all atmospheric pressure levels compared to standard deviations ranging from 3.5° to 7.5°C. For the complete data set, the standard error of estimate averages about 1°C greater than for clear skies only. It should be pointed out that part of the

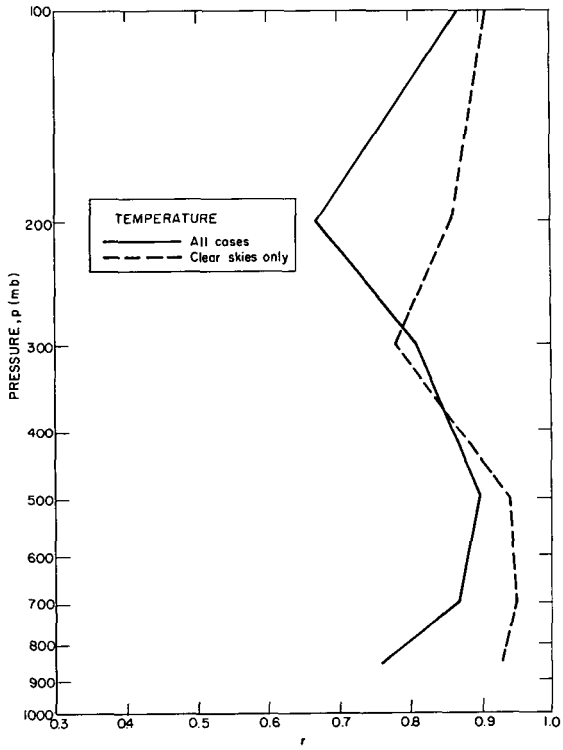

FIGURE 2.—Coefficient of multiple correlation (r) between temperature and infrared spectral predictors for the complete data set and for clear skies only.

standard error of estimate is a result of errors in the meteorological data that are used to develop the regression relations. Although confined to the 30°–60°N latitude sector, much of the meteorological data are from oceanic regions, and lack of radiosonde observations in these regions causes uncertainties in the meteorological analyses. There is no doubt that lower errors would result if the regression relationships were developed and verified against more accurate meteorological data. For example, Smith et al. (1970) obtain a root-mean-square (rms) temperature error of 1.5°–2.0°C for clear skies when

they verify their statistical technique against data over Western Europe where the satellite data are within 3 hr of the standard radiosonde release times.

It is of interest to compare the results for the complete data set with the results obtained by Smith et al. (1970) for cloudy cases. In their technique, they use an iterative procedure to eliminate the effect of clouds. The procedure requires the use of a measured or estimated surface temperature to determine whether clouds are present or not. The surface temperature is compared to the brightness temperature in the window wavelength of the satellite infrared spectrometer (SIRS) and if it is more than 5°C higher than the observed window brightness temperature, clouds are assumed to be present. For such cloudy conditions, Smith et al. (1970) obtain rms temperature errors that average 2.0°–2.5°C when verified against reasonably accurate meteorological data. This compares to an average standard error of estimate of about 3.5°C in the present study. However, the present regression equation for the complete data set does not require any information on surface temperature—only satellite observations are used to specify the temperature structure. As a result, the errors are somewhat larger than in the case where surface temperature is known and corrections for clouds are made.

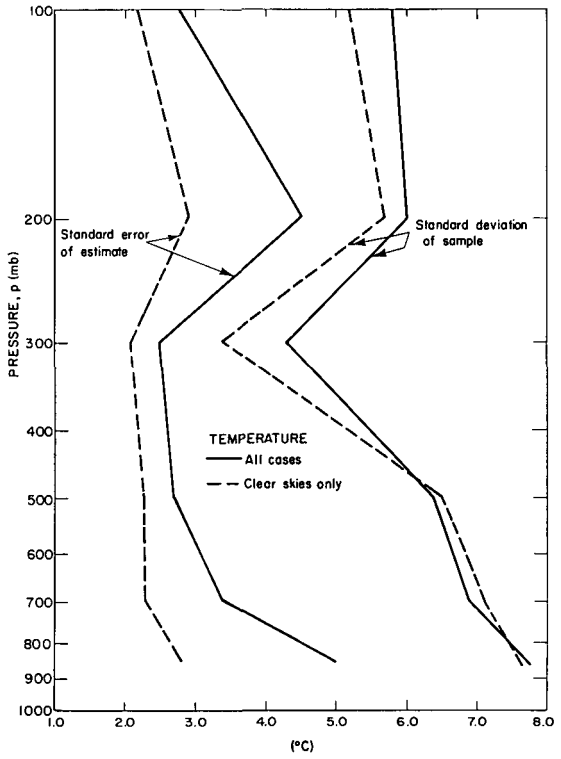

FIGURE 3.—Standard errors of estimate (°C) and standard deviations (°C) for temperature for the complete data set and for clear skies only.

The standard error estimates obtained for the complete data set probably are good estimates of the maximum errors that would result from operational application of the stepwise regression technique to the IRIS spectra. Separation of cloudy from clear conditions (by means of either surface temperature information or satellite cloud pictures) and development and verification of the regression relationships on reasonably accurate meteorological data should bring temperature errors down to the 1°–2°C range. The global meteorological network is striving for ±1°C temperature accuracy. However, where conventional observations are not available, even 2°C accuracy could be useful.

Geopotential Heights

According to the hypsometric equation, the geopotential height of a pressure level depends upon the pressure at the earth's surface and the mean temperature of the

TABLE 6.—*Results of stepwise multiple regression analysis of geopotential height* (m) using complete data set*

Pressure level (mb)	Coefficient of multiple correlation	Predictors		Intercept and regression coefficients		
850	0. 80	T_B(min)		6, 260	−28. 2	
700	. 86	T_B(min)		9, 148	−41. 2	
500	. 89	T_B(min)		14, 091	−63. 4	
300	. 91	T_B(min)		20, 634	−92. 8	
200	. 92	T_B(min)	712	18, 459	−91. 2	7. 60
100	. 88	T_B(min)	$\nu[T_B(\text{min})]$	17, 732	−57. 0	−7. 36

*Departures from a constant value at each pressure level

TABLE 7.—*Results of stepwise multiple regression analysis of geopotential height* (m) for clear-sky data*

Pressure level (mb)	Coefficient of multiple correlation	Predictors			Intercept and regression coefficients			
850	0. 88	702	450		9, 636	−48. 8	5. 44	
700	. 92	702	744		8, 063	−51. 8	14. 2	
500	. 94	691	702	744	10, 061	−6. 68	−63. 2	22. 4
300	. 94	1198	702	733	10, 636	5. 85	−86. 6	29. 0
200	. 94	1198	691		5, 005	15. 2	−41. 6	
100	. 94	1064	691	1052	2, 176	24. 7	−21. 4	−15. 6

*Departures from a constant value at each pressure level

atmosphere between the earth's surface and the pressure level. Thus, one would expect the geopotential height of a pressure level to be correlated with the mean temperature of the layer below the pressure level. Such correlations have indeed been found by Smith and Fritz (1969) who also find that upper tropospheric heights are also highly correlated with lower stratospheric temperatures. These correlations between the temperature and height field suggest that the infrared spectral data, being sensitive to the temperature field, can be used to specify the geopotential height field. Exploiting this possibility, Smith et al. (1970) have successfully obtained geopotential heights from the SIRS experiment on Nimbus 3.

Results of the stepwise multiple regression analyses with geopotential height as a predictand are shown in table 6 for the complete data set and in table 7 for clear skies only. As in the specification of temperature, T_B(min) is the best predictor when all 500 spectra are included. For clear skies, predictors are mainly brightness temperatures in the CO_2 band, although at 300 and 200 mb, a brightness temperature in the 6.3-μm H_2O band is the leading predictor, and at 100 mb a brightness temperature in the 9.6-μm O_3 band is the leading predictor.

A comparison of the multiple correlation coefficient for the complete data set and for clear skies only is presented in figure 4. For clear skies, all but the lowest level (850 mb) have multiple correlations greater than 0.9; for the complete data set, all but the lowest level have multiple correlations greater than 0.85. The correlation coefficients are greater for geopotential height than for temperature. Presumably, this is due to the dependence of geopotential height of a pressure surface on the mean temperature of an atmospheric layer, which, because of broad weighting functions, is what the infrared intensities are sensitive to. Or put another way, if the errors of the satellite-specified temperatures at the different pressure levels are random with respect to each other, then the mean temperature of a number of levels (and, hence, the geopotential height) can be specified more accurately then the temperature at a single level.

The standard errors of estimate and standard deviations for geopotential height are plotted in figure 5. For clear skies, the standard error of estimate has an average value of about 70 m; for the complete data set, the standard error of estimate has an average of about 85 m. As in the case of temperature, parts of these errors are due to errors in the meteorological analyses and could be reduced by developing and verifying the regression equations with accurate meteorological data. The accuracy of geopotential heights obtained from radiosonde observations is about 0.4 percent (Leviton 1971), and this value represents a goal for any satellite technique. At 500 mb, this means an rms error of 20 m; at 100 mb, the rms error is 64 m. At present, the main contribution of satellite-derived geopotential heights would be in regions of the world that lack radiosonde observations. Such application of the statistical technique of Smith et al. (1970) has been made. For clear skies, they obtain rms errors of approximately 30 m, and for cloudy skies, about 50–60 m. However, their technique of height

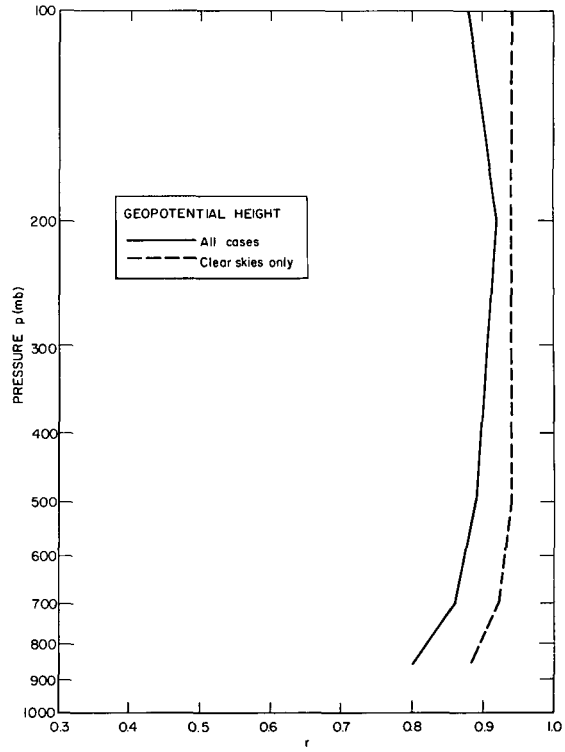

FIGURE 4.—Coefficient of multiple correlation (r) between geopotential height and infrared spectral predictions for the complete data set and for clear skies only.

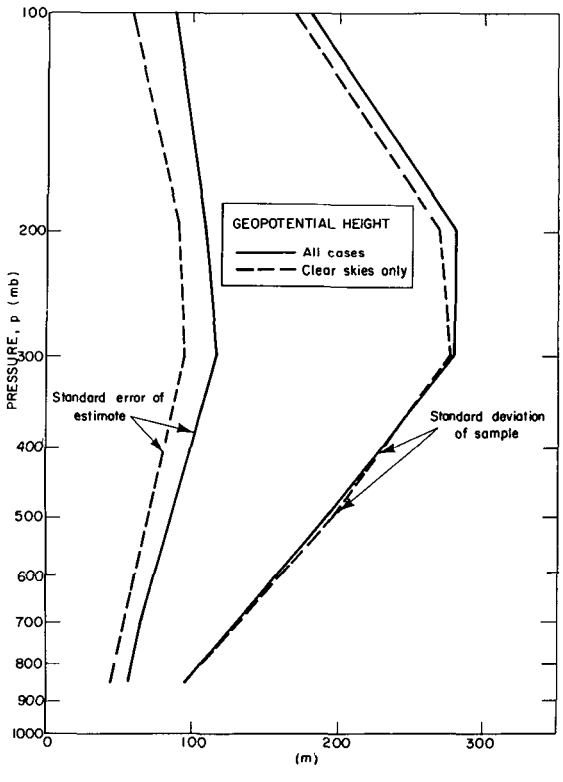

FIGURE 5.—Standard errors of estimate (m) and standard deviations (m) for geopotential height for the complete data set and for clear skies only.

specification makes use of the observed 850-mb geopotential height to correct the SIRS statistically derived heights.

One method of reducing the errors of the regression technique is to average the specifications obtained from

TABLE 8.—*Results of stepwise multiple regression analysis of tropopause (trop.) characteristics using the complete data set*

Parameter	Coefficient of multiple correlation	Predictors		Intercept and regression coefficients		
Trop. pressure (cb)	0. 79	691	$\nu[T_B(\min)]$	−158	1. 37	−0. 177
Trop. temperature (°C)	0. 58	691	$\nu[T_B(\min)]$	49. 5	0. 859	−0. 429

several consecutive satellite spectra. Consecutive spectra are of the order of 100 km apart in the horizontal. Global meteorological observational requirements for the Global Atmospheric Research Program (Tepper and Ruttenberg 1970) call for a horizontal resolution of 400 km. Thus, averaging of the specifications of three or four consecutive spectra would reduce errors while maintaining the horizontal resolution within required limits.

Tropopause Characteristics

Results of the stepwise multiple regression analyses with tropopause pressure and tropopause temperature as predictands are shown in table 8. As discussed in section 2, we expected that T_B (min) would be a good predictor of tropopause temperature, and that $\nu[T_B$ (min)], the wave number of minimum brightness temperature along the side of the 15-μm band, would be a good predictor of tropopause altitude (or pressure). The results shown in table 8 do indicate that $\nu[T_B(\min)]$ is a predictor for specification of tropopause pressure. However, $T_B(\min)$ is not included as a predictor of tropopause temperature. Somewhat puzzled, we examined the correlation matrix for the specification of tropopause characteristics and found that the best single predictor of tropopause temperature is $T_B(691\ \text{cm}^{-1})$ with a correlation coefficient of 0.43. However, $T_B(\min)$ is a close second with $r=0.41$; $\nu[T_B(\min)]$ is third with $r=-0.40$. In the stepwise multiple regression analysis for tropopause temperature, $T_B(691\ \text{cm}^{-1})$ is therefore selected as the first predictor. The selection of $\nu[T_B(\min)]$ rather than $T_B(\min)$ as the second predictor is due to $T_B(\min)$ being highly correlated with the first predictor, $T_B(691\ \text{cm}^{-1})$, whereas $\nu[T_B(\min)]$ is not. These correlations among the different brightness temperatures indicate redundancies in the observed spectra that, as pointed out previously, can be exploited by deriving two or more regression relationships and averaging the results of each equation to obtain an estimate of the predictand.

The multiple correlation coefficients indicate that tropopause pressure ($r=0.79$) can be more accurately specified than tropopause temperature ($r=0.58$). A comparsion of the standard error estimates with the standard deviations in the complete data set is presented in table 9. These results indicate that tropopause pressure can be specified to within ± 32 mb and tropopause temperature to ± 4.8°C from satellite infrared spectra. It is not immediately evident whether the specification of tropopause pressure to this accuracy is of operational usefulness.

TABLE 9.—*Standard deviations of sample and standard errors of estimate for tropopause characteristics*

Parameter	Standard deviation of sample	Standard error of estimate
Tropopause pressure (mb)	52	32
Tropopause temperature (°C)	5. 9	4. 8

4. CONCLUSIONS

Exploratory studies with Nimbus 3 IRIS data indicate that, in addition to temperature, such meteorological parameters as geopotential heights of pressure surfaces, tropopause pressure, and tropopause temperature can be inferred from the observed spectra with the use of simple regression equations. The technique of screening the IRIS spectral data by means of stepwise regression to obtain the best radiation predictors of meteorological parameters is validated. The simplicity of application of the technique and the simplicity of the derived linear regression equations—which contain only a few terms—suggest usefulness for this approach. The redundancy found in the complete spectrum—brightness temperatures at one wave number may be just as well correlated with a particular meteorological parameter as brightness temperatures at other wave numbers—suggests the possibility of deriving a set of regression equations for the specfication of a meteorological parameter, each equation having different predictors. By averaging the results of these equations, one can eliminate some of the random instrumental errors associated with the observed spectrum. The results obtained for the inference of geopotential heights, tropopause pressure, and tropopause temperature confirm the hypothesis that meteorological parameters that are related to the primary variables influencing the observed infrared spectrum can also be inferred directly from the spectrum.

It was not the purpose of this paper to analyze in detail the causes of the observed correlations between the brightness temperatures and the meteorological parameters. Part of the observed correlation may be due to latitudinal trends in the predictors and in the meteorological parameters. For example, part of the correlation of T_B (min), which is representative of temperatures in the tropopause region, with lower tropospheric temperatures may be due to the latitudinal trends of both tropopause temperature and lower atmospheric temperatures. On the other hand, part of the correlation may occur because the variations of tropopause temperature and lower atmospheric temper-

ature around a latitude circle (or in time) are correlated with each other at any particular latitude. The existence of such correlations between the temporal variations of tropopause temperature and lower atmosphere temperatures at individual weather stations have been known for some time. (See e.g., Haurwitz 1941, pp. 320–327, who reviews the 1919 work of Dines and others.) It would be of interest to determine whether the correlations and standard errors of estimate would be greater or smaller if one were to perform a similar analysis for a single latitude or for individual stations over a long time period. Such a comparison of latitudinal versus temporal effects can only be carried out with a larger sample of data than that used in this exploratory study and is strongly recommended as a follow-on to the present study. It is also suggested that verification tests be performed on independent data samples and that the regression relationships be derived and verified for regions where conventional meteorological data are plentiful. Additional recommendations for further research along these lines are as follows:

1. The feasibility of inferring other meteorological parameters—such as winds, clouds, water vapor, and jet stream characteristics—directly from the IRIS spectra should be investigated.
2. The spectral and spatial redundancy of the IRIS data should be exploited by deriving sets of regression relationships for a single meteorological parameter and by combining results for several consecutive spectra.
3. Techniques for improving meteorological inferences in the presence of cloudy skies should be developed.
4. The possibility should be explored of using stepwise regression analysis techniques on the IRIS spectra for the purpose of determining which wavelengths should be observed by an N channel radiometer for atmospheric inferences from an operational meteorological satellite.

ACKNOWLEDGMENTS

The author thanks Barney Conrath and Vince Salomonson for providing the IRIS data and conventional meteorological data used in this study and for helpful information concerning the data.

This research was supported by the NASA Goddard Space Flight Center under Contract No. NA5-21024.

REFERENCES

Bennett, C. A., and Franklin, N. L., *Statistical Analysis in Chemistry and the Chemical Industry*, John Wiley & Sons, Inc., New York, N.Y., **1954**, 724 pp.

Chekirda, A. Z., and Iakovleva, T. A., "Ispol'zovanie Sputnikovoĭ Informatsii ob Ukhodiashchem Izluchenii v Tseliakh Opredeleniia Zon Vypadeniia Osadkov" (Use of Satellite Information of Outgoing Radiation to Determine Precipitation Zones), *Meteorologiia i Gidrologiia* No. 6, Hydrometeorological Service of the U.S.S.R., Moscow, June **1967**, pp. 13–20.

Conrath, Barney J., "On the Estimation of Relative Humidity Profiles From Medium-Resolution Infrared Spectra Obtained From a Satellite," *Journal of Geophysical Research, Oceans and Atmospheres*, Vol. 74, No. 13, June 20, **1969**, pp. 3347–3361.

Conrath, Barney J., Hanel, R. A., Kunde, V. G., and Prabhakara, C., "The Infrared Interferometer Experiment on Nimbus 3," *Journal of Geophysical Research, Oceans and Atmospheres*, Vol. 75, No. 30, Oct. 20, **1970**, pp. 5831–5857.

Goddard Space Flight Center, NASA, *Nimbus III User's Guide*, National Aeronautics and Space Administration, Greenbelt, Md., **1969**, 238 pp.

Hanel, R. A., Schlachman, B., Clark, F. O., Prokesh, C. H., Taylor, J. B., Wilson, W. M., and Chaney, L., " The Nimbus III Michelson Interferometer," *Applied Optics*, Vol. 9, No. 8, Aug. **1970**, pp.1767–1774.

Haurwitz, Bernhard, *Dynamic Meteorology*, 1st Ed., McGraw-Hill Book Co., Inc., New York, N.Y., **1941**, 365 pp.

Jensen, Clayton E., Winston, Jay S., and Taylor, V. Ray, "500-Mb. Heights as a linear Function of Satellite Infrared Radiation Data," *Monthly Weather Review*, Vol. 94, No. 11, Nov. **1966**, pp. 641–649.

Leviton, Robert, Mar. **1971** (personal communication).

Martin, Frank L., and Salomonson, Vincent V., "Statistical Characteristics of Subtropical Jet-Stream Features in Terms of MRIR Observation From Nimbus II," *Journal of Applied Meteorology*, Vol. 9, No. 3, June **1970**; pp. 508–520.

Rodgers, C.D., "The Inversion Problem in the Presence of Cloud," *World Meteorological Organization Technical Note* No. 104, Geneva, Switzerland, **1970**, pp. 73–77.

Shenk, William E., "TIROS II Window Radiation and Large Scale Vertical Motion," *Journal of Applied Meteorology*, Vol. 2, No. 6, Dec. **1963**, pp. 770–775.

Smith, William L., and Fritz, S., "On the Statistical Relation Between Geopotential Height and Temperature-Pressure Profiles," *ESSA Technical Memorandum* NESCTM–18, U.S. Department of Commerce, National Environmental Satellite Center, Washington, D.C., Nov. **1969**, 16 pp.

Smith, William L., Woolf, H. M., and Jacob, W. J., "A Regression Method for Obtaining Real-Time Temperature and Geopotential Height Profiles From Satellite Spectrometer Measurements and Its Application to NIMBUS 3 'SIRS' Observations," *Monthly Weather Review*, Vol. 98, No. 8, Aug. **1970**, pp. 582–603.

Tepper, Morris, and Ruttenberg, Stanley, "The Role of the Satellite in Future Observing Systems," *Meteorological Monographs*, Vol. 11, No. 33, American Meteorological Society, Boston, Mass., Oct. **1970**, pp. 56–70.

Twomey, S., "The Application of Numerical Filtering to the Solution of Integral Equations Encountered in Indirect Sensing Measurements," *Journal of the Franklin Institute*, Vol. 279, No. 2, Pergamon Press, New York, N.Y., Feb. **1965**, pp. 95–109.

Twomey, S., "Indirect Measurements of Atmospheric Temperature Profiles From Satellites: II. Mathematical Aspects of the Inversion Problem," *Monthly Weather Review*, Vol. 94, No. 6, June **1966**, pp. 363–366.

Westwater, Ed R., and Strand, Otto N., "Statistical Information Content of Radiation Measurements Used in Indirect Sensing," *Journal of the Atmospheric Sciences*, Vol. 25, No. 5, Sept. **1968**, pp. 750–758.

Zak, J. Allen, and Panofsky, Hans A., "Estimation of Stratospheric Flow From Satellite 15-Micron Radiation," *Journal of Applied Meteorology*, Vol. 7, No. 1, Feb. **1968**, pp. 136–140.

Zhvalev, V. F., Kondratyev, K., and Ter-Markaryants, N., "The Use of Infrared Photographs of the Earth to Trace the Dynamics of Ocean Currents and to Identify Jet Streams and Noctilucent Clouds," *Atmospheric and Oceanic Physics*, Vol. 3, No. 11, The American Geophysical Union, Washington, D.C., Nov. **1967**, pp. 699–703 (translation of O vozmozhnosti ispol'zovaniia infrakrasnykh izobrazhenii Zemli diia proslezhivaniia dinamiki morskikih techenii, obnaruzheniia struinykh techenii i serebristylkh oblakov, *Akademiia Nauk SSSR, Isvestiia, Fizika Atmosfery i Okeana*, Vol. 3, No. 11, Leningrad, U.S.S.R., Nov. 1967, pp. 1187–1195)

[Received April 30, 1971; revised November 17, 1971]